

Arctic Oscillation response to volcanic eruptions in the IPCC AR4 climate models

Georgiy Stenchikov,¹ Kevin Hamilton,² Ronald J. Stouffer,³ Alan Robock,¹
V. Ramaswamy,³ Ben Santer,⁴ and Hans-F. Graf⁵

Received 29 May 2005; revised 4 January 2006; accepted 23 January 2006; published 11 April 2006.

[1] Stratospheric sulfate aerosol particles from strong volcanic eruptions produce significant transient cooling of the troposphere and warming of the lower stratosphere. The radiative impact of volcanic aerosols also produces a response that generally includes an anomalously positive phase of the Arctic Oscillation (AO) that is most pronounced in the boreal winter. The main atmospheric thermal and dynamical effects of eruptions typical of the past century persist for about two years after each eruption. In this paper we evaluate the volcanic responses in simulations produced by seven of the climate models included in the model intercomparison conducted as part of the preparation of the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). We consider global effects as well as the regional circulation effects in the extratropical Northern Hemisphere focusing on the AO responses forced by volcanic eruptions. Specifically we analyze results from the IPCC historical runs that simulate the evolution of the circulation over the last part of the 19th century and the entire 20th century using a realistic time series of atmospheric composition (greenhouse gases and aerosols). In particular, composite anomalies over the two boreal winters following each of the nine largest low-latitude eruptions during the period 1860–1999 are computed for various tropospheric and stratospheric fields. These are compared when possible with observational data. The seven IPCC models we analyzed use similar assumptions about the amount of volcanic aerosols formed in the lower stratosphere following the volcanic eruptions that have occurred since 1860. All models produce tropospheric cooling and stratospheric warming as in observations. However, they display a considerable range of dynamic responses to volcanic aerosols. Nevertheless, some general conclusions can be drawn. The IPCC models tend to simulate a positive phase of the Arctic Oscillation in response to volcanic forcing similar to that typically observed. However, the associated dynamic perturbations and winter surface warming over Northern Europe and Asia in the post-volcano winters is much weaker in the models than in observations. The AR4 models also underestimate the variability and long-term trend of the AO. This deficiency affects high-latitude model predictions and may have a similar origin. This analysis allows us to better evaluate volcanic impacts in up-to-date climate models and to better quantify the model Arctic Oscillation sensitivity to external forcing. This potentially could lead to improving model climate predictions in the extratropical latitudes of the Northern Hemisphere.

Citation: Stenchikov, G., K. Hamilton, R. J. Stouffer, A. Robock, V. Ramaswamy, B. Santer, and H.-F. Graf (2006), Arctic Oscillation response to volcanic eruptions in the IPCC AR4 climate models, *J. Geophys. Res.*, *111*, D07107, doi:10.1029/2005JD006286.

¹Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey, USA.

²International Pacific Research Center, University of Hawaii, Honolulu, Hawaii, USA.

³NOAA Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, New Jersey, USA.

⁴Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, California, USA.

⁵Department of Geography, University of Cambridge, Cambridge, U.K.

1. Introduction

[2] Strong explosive volcanic eruptions are known to have a large, if transient, impact on climate [Robock, 2000]. Typically after a major explosive eruption a global-mean surface cooling is observed, which is an expected consequence of the increased reflection of solar radiation by long-lived aerosols in the stratosphere [Humphreys, 1913, 1940; Mitchell, 1961]. Also observed is a warming of the equatorial lower stratosphere attributable to enhanced absorption of terrestrial IR and solar near-IR radiation [Labitzke *et al.*,

1983; Quiroz, 1983; Parker and Brownscombe, 1983; Labitzke and McCormick, 1992; Labitzke, 1994; Stenchikov et al., 1998]. In addition, stratospheric aerosols serve as surfaces for heterogeneous reactions liberating anthropogenic chlorine accumulated in the stratosphere in the late 20th century, producing ozone depletion. Observations and model simulations also show that strong low-latitude eruptions are generally followed by an enhanced positive phase of the Arctic Oscillation (AO) for one or two Northern Hemisphere (NH) winters [Groisman, 1992; Robock and Mao, 1992, 1995; Graf et al., 1993; Kirchner et al., 1999; Stenchikov et al., 2002, 2004].

[3] The radiative perturbation caused by stratospheric aerosols produced by a major volcano provides a possible test of the ability of general circulation models (GCMs) to respond realistically to global-scale radiative forcing [Stenchikov et al., 2002; Robock, 2003]. There have been several earlier studies in which the radiative perturbations associated with a major volcano have been imposed in a GCM and the nature of the simulated temperature and circulation responses investigated. Many of these studies have focused on simulating the aftermath of the Mount Pinatubo eruption, which was both the largest eruption of the 20th century and the eruption for which the stratospheric aerosol has been best observed. There have been two main foci of such studies, analysis of the simulation of global-mean surface temperature response and simulation of the response of the extratropical circulation in the NH winter season. With regard to the NH winter circulation, it is noteworthy that the observed long-term trends in the last few decades include a component that is consistent with a significant increase in the index of the AO [Hurrell, 1995; Thompson and Wallace, 1998; Ostermeier and Wallace, 2003]. This observed trend in the AO index is generally not well reproduced by current GCMs when forced with the historical trends of greenhouse gas and aerosol concentration [Osborn, 2004; Gillett, 2005; Knutson et al., 2006; Scaife et al., 2005].

[4] The use of volcanic simulations as tests of model climate feedbacks and sensitivity is unfortunately somewhat hampered by the limited observational data available: over the last 150 years major eruptions with expected global climate effects have occurred less than once per decade, on average. Any climate anomalies observed in the aftermath of these eruptions will also reflect other internally-generated variability (e.g., El Niño/Southern Oscillation (ENSO), quasi-biennial oscillation (QBO), chaotic weather changes) in the atmosphere-ocean system. With model simulations, one can perform multiple realizations to clearly isolate the volcanic climate signal, but the real world data are limited to the single realization during the period since quasi-global instrumental records have been available.

[5] Despite the limited observations it is still a valuable exercise to examine the response of GCMs to historical volcanic eruptions and compare with the data available. This was one motivation for several GCM groups to include stratospheric aerosols from observed eruptions in their specification of the evolution of atmospheric composition in the “20th century” integrations conducted as part of the model intercomparison for the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). The present paper reports on an analysis of the results from

a total of seven of the models in the AR4 intercomparison. The analysis here is focused on extracting the typical or average response to volcanic aerosol loading in the various models. To provide as robust statistics as possible and allowing results for several models to be summarized concisely, much of the analysis involves computation of composites of the circulation anomalies in the simulations over the first two years after each of the nine largest low-latitude volcanic eruptions in the post-1860 era.

[6] This paper is organized as follows. Section 2 reviews the relevant background, including previous GCM studies of volcanic effects. Section 3 introduces the GCMs considered in the present study and discusses the design of the IPCC model integrations that have been analyzed here. Section 4 considers the results obtained for the various models, including both global and regional effects that can be attributed to the inclusion of volcanic aerosols. Results and conclusions are summarized in section 5.

2. Background and Review of Earlier Model Studies

[7] Volcanic eruptions, like the one of Mount Pinatubo in 1991, which had a global visible optical depth maximizing at about 0.15, cause negative perturbation of the 60°S–60°N averaged radiative balance at the top of the atmosphere reaching -2.5 – 3.0 W/m² and global surface cooling of about -0.5 K. The 1991 Mount Pinatubo eruption was the strongest and best observed in the 20th century. Therefore it is not surprising that most previous studies on volcano/AO interactions were conducted so far for this eruption [e.g., Stenchikov et al., 1998; Kirchner et al., 1999; Ramachandran et al., 2000; Stenchikov et al., 2002, 2004; Yang and Schlesinger, 2002; Shindell et al., 2004].

[8] Low-latitude volcanic eruptions force a positive phase of the AO (associated with stronger westerlies and winter warming over Northern Eurasia and North America) because of aerosol radiative heating in the equatorial lower stratosphere that strengthens the equator-to-pole temperature gradient in the lower stratosphere and accelerating the polar vortex [Kodera, 1994; Perlwitz and Graf, 1995; Ohhashi and Yamazaki, 1999; Kirchner et al., 1999; Kodera and Kuroda, 2000a, 2000b; Shindell et al., 2001]. The strengthening of the polar jet is amplified by a positive feedback between the polar NH winter vortex and vertical propagation of planetary waves. The stronger vortex reflects planetary waves decreasing deceleration and preserving axial symmetry of the flow. Stenchikov et al. [2002] also found that tropospheric cooling caused by volcanic aerosols can affect storminess and generation of planetary waves in the troposphere. This tends to decrease the flux of wave activity and negative angular momentum from the troposphere into the polar stratosphere reducing wave drag on the vortex. Polar ozone depletion caused by heterogeneous chemistry initiated by volcanic aerosols in the lower stratosphere tends to cool the polar stratosphere in spring, strengthening the polar vortex and delaying final warming [Stenchikov et al., 2002].

[9] With respect to the dynamical mechanisms through which perturbations of the stratospheric annular circulation in general can influence tropospheric annular modes, Song and Robinson [2004] pointed out that tropospheric wester-

Table 1. IPCC Models and Their Treatment of Volcanic Aerosols^a

Mark	Model Name	Spatial Resolution	Model Top, hPa	Volcanic Aerosols	Beginning of Run	Ensemble Members
a	GFDL CM 2.0	2° × 2.5° L24	3.0	<i>Sato et al.</i> [1993], <i>Stenchikov et al.</i> [1998]	1861	3
b	GFDL CM2.1	2° × 2.5° L24	3.0	<i>Sato et al.</i> [1993], <i>Stenchikov et al.</i> [1998]	1861	5
c	GISS-EH	4° × 5° L20	0.1	<i>Sato et al.</i> [1993]	1880	5
d	GISS-ER	4° × 5° L20	0.1	<i>Sato et al.</i> [1993]	1880	9
e	NCAR CCSM3	T85 L26	2.2	<i>Ammann et al.</i> [2003]	1870	6
f	NCAR PCM1	T42 L18	2.9	<i>Ammann et al.</i> [2003]	1890	4
g	MIROC-medres	T42 L20	10.0	<i>Sato et al.</i> [1993]	1850	3
	MIROC-hires	T106 L56	1.0	<i>Sato et al.</i> [1993]	1900	1
	MRI	T42 L30	0.4	Modified solar constant	1851	5

^aThe models are as follows: GFDL CM2.0 and CM2.1, two versions of the GFDL coupled model with different atmospheric dynamical core [*Delworth et al.*, 2006; *Gnanadesikan et al.*, 2006]; GISS EH and ER, two versions of the GISS coupled model with different ocean modules (<http://www.giss.nasa.gov/tools/modelE>); NCAR CCM3, Community Climate System Model, version 3.0 (<http://www.cesm.ucar.edu>); NCAR PCM1, Parallel Climate Model [*Kiehl et al.*, 1998]; MIROC medres and hires, developed jointly at the Center for Climate System Research, University of Tokyo, National Institute for Environmental Studies, Frontier Research Center for Global Change, and Japan Agency for Marine-Earth Science and Technology (<http://www.ccsr.u-tokyo.ac.jp/kyosei/hasumi/MIROC/tech-repo.pdf>); Meteorological Research Institute (MRI), Japan [*Yukimoto et al.*, 2001].

lies can be strengthened by changes of planetary wave vertical propagation and/or reflection within the stratosphere and associated wave-zonal flow interaction [*Kodera*, 1994; *Perlwitz and Graf*, 1995; *Perlwitz and Harnik*, 2003], downward control, or the nonlinear effect of baroclinic eddies [*Black*, 2002; *Black and McDaniel*, 2004; *Haynes et al.*, 1991; *Limpasuvan et al.*, 2004]. All these mechanisms could play a role in shaping tropospheric dynamic response to volcanic forcing.

[10] The improved spatial resolution and physical parameterizations used in the up-to-date climate models do not always guarantee a correct description of stratosphere-troposphere dynamic coupling. *Shindell et al.* [2001] reported that numerical experiments with an old version of the Goddard Institute for Space Studies (GISS) 23-layer middle atmosphere model with the top at 85 km suggested that volcanic forcing caused a positive phase of the AO. *Shindell et al.* [2004] used an early version of the new GISS ModelE with the top at 0.1 hPa, and found that it produced better AO responses to volcanic forcing than the older middle atmosphere model. They estimated the climate responses during the cold season to the largest volcanic eruptions since 1600 using instrumental data and proxy-based reconstruction. They showed that because of internal climate variability, climate responses from individual eruption are not representative and hence conducted a composite analysis for the largest volcanic eruptions demonstrating a statistically significant winter warming pattern. They claimed that a GCM has to vertically well resolve processes in the middle atmosphere to be able to reproduce this effect.

[11] *Broccoli et al.* [2003] conducted a series of historic simulations from 1865 to 1997 implementing precalculated volcanic forcing of *Andronova et al.* [1999] in the Geophysical Fluid Dynamics Laboratory (GFDL) R30 14-layer GCM coupled with the global GFDL Modular Ocean Model MOM 1.1. Their results imply that because of significant differences in the models' basic climate states and radiative transfer schemes it is preferable to calculate the aerosol radiative impact interactively by implementing aerosol optical properties and distribution directly into a model.

[12] *Oman et al.* [2005] used the GISS ModelE to simulate the climate impact of the 1912 Katmai eruption in Alaska. They calculated 20-member ensembles of simulations and found that a volcanic aerosol cloud which

remained mostly north of 30°N could not produce a significant winter warming pattern for a hemispheric optical depth higher than for the Pinatubo eruption in 1991. This is because in the winter the heating of the mid-latitude volcanic cloud is too weak to produce a sufficient meridional temperature gradient in the lower stratosphere and to influence the polar vortex. With a similar amount of aerosol produced by the midlatitude Laacher See eruption, *Graf and Timmreck* [2001] found a strengthened polar vortex in the ECHAM4 model due to increased longwave cooling of the dense aerosol layer over the pole in winter, but conducted only a single realization of their GCM experiment.

3. Models and Experiments

[13] As part of the IPCC intercomparison for the AR4, model groups were encouraged to perform historical “20th century” integrations. These generally started from the late 19th century and proceeded through 1999 or 2000. In these integrations, a detailed time series of atmospheric composition (long-lived greenhouse gases and atmospheric aerosols) was specified based on available observations. However, there was no standardization of the atmospheric composition time-series imposed, and each group adopted somewhat different approaches. For the present study we examined the data from the “20th century” integrations from the 19 coupled ocean-atmosphere GCMs for which data had been deposited in the IPCC data archive by the end of February 2005. Of these models only nine included some treatment of the effects of volcanic eruptions. Some basic information about these models is summarized in Table 1. Of these we excluded two models from further analysis, MRI because the volcanic effects were imposed simply as a reduction in the solar constant, rather than adopting a more realistic treatment, and MIROC-hires because it had only a single realization and this only spanned the period starting in 1900. For each of the seven models (from four different scientific centers) selected for further analysis, there are at least three realizations (with the same forcing but different initial conditions) of the historical simulations.

[14] Each model group adopted their own specification of the volcanic aerosols imposed in their runs. In each case a time-height specification of the zonal-mean aerosol concentrations and properties was constructed. The aerosol data set

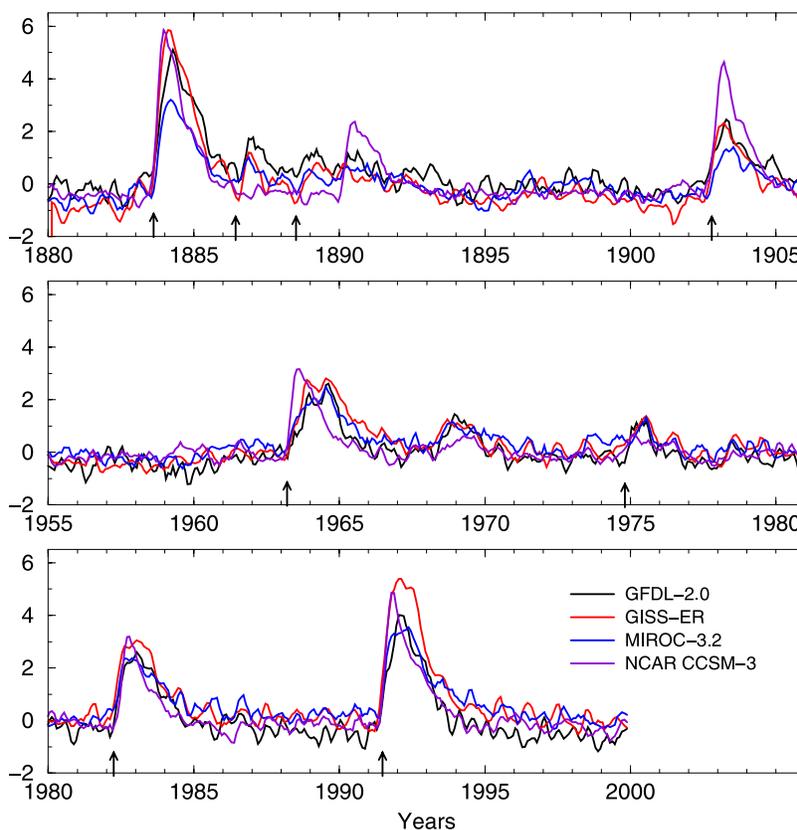


Figure 1. Three-month running means of the anomalies in global-averaged reflected solar flux (W m^{-2}) for selected periods. Results plotted for the four models from each modeling group (Table 1) and represent averages over all the individual realizations for each model. The arrows show the times of the major eruptions listed in Table 2.

that served as the basis for most model groups is that of *Sato et al.* [1993], improved by *Hansen et al.* [2002], which provided zonal-mean vertically resolved aerosol optical depth for visible wavelengths and column average effective radii. The two National Center for Atmospheric Research (NCAR) models in the group (CCM3 and PCM) based their specification of volcanic aerosol on a data set of *Ammann et al.* [2003] who calculated the zonal-mean aerosol spatial distribution using estimated aerosol loadings and a diffusion-type parameterized transport algorithm assuming fixed effective radius of $0.42 \mu\text{m}$ for determining aerosol optical properties.

[15] For the GFDL CM2.0 and CM2.1 models, the aerosol effective radii of *Hansen et al.* [2002] were modified using Upper Atmosphere Research Satellite observations for the Pinatubo period, accounting for their variations with altitude, especially at the top of the aerosol layer where particles became very small. Then aerosol optical characteristics were calculated following *Stenchikov et al.* [1998] using optical depths from *Sato et al.* [1993] and *Hansen et al.* [2002].

[16] Unfortunately, the model groups did not provide the quantitative radiative forcing associated with the composition changes imposed during the historical runs. Such values would have had to have been diagnosed from detailed calculations with the model radiative transfer schemes. From the data provided in the IPCC archive, the

closest we can come to diagnosing the global-mean radiative forcing of volcanic aerosols is through examination of the time series of the global-mean reflected top-of-the-atmosphere (TOA) solar radiation. This time series was averaged for all the realizations performed for each model. The GFDL, GISS, and NCAR pairs of models (Table 1) share radiative schemes and volcanic aerosol input characteristics and therefore behaved similarly. Therefore in Figure 1 we show results only for one model from each modeling group. Only selected segments of the 1880–1999 period are shown, but these include eight of the nine largest low-latitude eruptions during the period. This diagnosed quantity includes the effects on the radiative balance from the aerosols that is generally considered part of the climate forcing, but also includes a contribution from the changes in albedo that are part of the response to the forcing (e.g., due to changes in clouds or snow cover). With this caveat, the curves in Figure 1 in post-volcanic periods should provide a rough comparison of the overall forcing of the global-mean thermal balance from the imposed aerosol in the different models. The increased reflectivity of the stratospheric aerosol is very evident in all the models after the major eruptions of Krakatau (1883), Santa María (1902), Agung (1963), El Chichón (1982) and Pinatubo (1991). Periods of more modest increases in reflectivity are also visible in at least some of the models after the eruptions of Tarawera (1886), Bandai (1888) and Fuego (1974). The volcanic

Table 2. Low-Latitude Volcanic Eruptions Chosen for Compositing^a

Volcano Name	Eruption Date	Latitude	Winters Analyzed	Reference Period	VEI	Optical Depth ($\lambda = 0.55 \mu\text{m}$) 30°S–30°N
Krakatau	Aug 27, 1883	6.10°S	1883–1884 1884–1885	1890–1901	6	0.20
Tarawera	Jun 10, 1886	38.23°S	1886–1887 1887–1888	1890–1901	5	0.07
Bandai	Jul 15, 1888	37.60°N	1888–1889 1889–1890	1890–1901	4	0.05
Santa María	Oct 24, 1902	14.76°N	1903–1904 1904–1905	1890–1901	6	0.10
Quizapu	Apr 10, 1932	35.65°S	1932–1933 1933–1934	1915–1931	5	0.02
Agung	Mar 17, 1963	8.34°S	1963–1964 1964–1965	1934–1955	4	0.11
Fuego	Oct 10, 1974	14.47°N	1975–1976 1976–1977	1965–1973	4	0.04
El Chichón	Apr 4, 1982	17.36°N	1982–1983 1983–1984	1976–1981	5	0.12
Pinatubo	Jun 15, 1991	15.13°N	1991–1992 1992–1993	1985–1990	6	0.18

^a“VEI” is the volcanic explosivity index [Newhall and Self, 1982]. Averaged over an equatorial belt, “optical depth” is calculated using the volcanic aerosol data set of Sato *et al.* [1993].

TOA effects are reasonably similar among the models, although the perturbations for the MIROC-medres model tend to be the smallest in each case, while the perturbations for the NCAR models tend to be the largest. There seems to be a tendency for the model results to agree more closely in the recent El Chichón and Pinatubo periods than in the periods after the other eruptions.

[17] For comparisons with observed surface air temperature, we use the HadCRUT2v data set (<http://www.cru.uea.ac.uk/cru/data/temperature>). This data set combines the land surface air temperature of Jones and Moberly [2003] with the HadSST1 SST data of Parker *et al.* [1995] and Rayner *et al.* [2003]. Variance adjustments have been applied to both land and ocean data [Jones *et al.*, 2001]. The HadSLP1 sea-level pressure data set for the period from 1871 to 1998 was obtained from http://hadobs.metoffice.com/gmslp/data/hadSLP1_1871-1998.asc.gz. It is an update of GMSLP2 using an improved land station data base, new interpolation scheme and the incorporation of local detail [Basnett and Parker, 1997]. We also used the National Centers for Environmental Prediction (NCEP)/NCAR reanalysis data for the second half of the 20th century [Kistler *et al.*, 2001].

[18] To provide concise measures of the volcanic effects that can be compared among several models, composites of the anomalies in the periods following the 9 largest low-latitude (40°S–40°N) eruptions since 1860 were made for each field of interest for each model. The locations and dates of the nine eruptions considered are given in Table 2. These are a subset of the volcanic events analyzed by Robock and Mao [1992]. We have excluded all high-latitude eruptions from the list of eruptions used by Robock and Mao because they appear to produce a qualitatively different effect on circulation than lower-latitude eruptions [Robock and Mao, 1995; Oman *et al.*, 2005]. We focus on the analysis of NH extratropical effects for a period consisting of the first two winters (December–February) following each of the eruptions that occurred not later in the year than August. For the two eruptions that occurred in October the two later winters were considered: 1903–1904 and 1904–1905 for Santa María, and 1975–1976 and 1976–1977 for Fuego.

[19] A complication in isolating the volcanic signal in either observations or the historical GCM runs is that there are significant long-term trends. So simply defining anomalies as the difference between the average during a particular post-volcanic period and the long-term mean is not appro-

priate. We have defined the anomalies in each post-volcano period relative to a reference period which is different for each eruption. The reference periods employed are given in Table 2, and in many cases were designed to represent the longest possible period immediately before the year of each eruption for which we can suppose the atmosphere was reasonably clear of volcanic aerosols. For the early Krakatau, Tarawera, Bandai and Santa María eruptions there were complications from the short intervals between eruptions and the fact that some model runs began only in 1880 or 1890. For these four eruptions anomalies are defined relative to a single 1890–1901 reference period. Anomalies appear to be fairly stable with respect to the choice of reference periods. For example, change of the reference period from 1861–1882 to 1890–1901 for Krakatau, Tarawera, and Bandai eruptions in the GFDL model runs did not make a sizable difference.

[20] A major complication in isolating the volcanic signal in the observed and model time series is the presence of signals from other sources of natural interannual variability. A particularly problematic issue is that a sampling of, say, nine two-year periods following the nine eruptions listed in Table 2 is likely to include contributions from an imperfectly sampled Southern Oscillation. We know that in the real world the period following both the El Chichón and Pinatubo eruptions coincided with ENSO events. For all the models we have at least three realizations, so to some extent we can hope to average out random Southern Oscillation signals by averaging over realizations. However, in practice there remain some sampling effects of this type in the IPCC runs. As an example, for the El Chichón and Pinatubo periods the GFDL CM2.0 and CM2.1 model simulations displayed average La Niña conditions (i.e., the opposite of that observed in the real world). The speculation that volcanic eruptions lead to a preference for El Niño conditions by Adams *et al.* [2003] and Mann *et al.* [2005] is not supported by these state-of-the-art GCM runs.

4. Results for Composited Volcanic Anomalies in Boreal Winter

[21] In this section we show results for each of the models in terms of anomaly fields composited over all realizations and over two winters for each of the nine volcanoes. Results are shown as maps and also as anomalies in a number of simple indices that we have defined. The indices are described briefly in the caption to Table 3 and the composite

Table 3. Integrated Model Responses^a

Model Name	Polar SLP, hPa	Atl. SLP, hPa	T _{GL} , K	T _{ES} , K	T _{AM} , K	T50, K	H50, m	U50/σ, m/s	Niño3.4, K
GFDL CM2.0	-0.13	0.6	-0.17	0.09	-0.70	0.50	-16.5	19.2/3.2	-0.18
GFDL CM2.1	-0.86	0.4	-0.12	0.25	-0.36	0.59	-41.8	20.9/2.2	-0.04
GISS-EH	-0.75	0.4	-0.10	0.06	-0.17	0.82	-29.3	19.9/2.6	-0.10
GISS-ER	-0.33	0.2	-0.06	-0.04	-0.06	0.74	-13.8	19.0/2.3	-0.08
NCAR CCSM3	0.10	0.4	-0.13	-0.32	-0.38	1.32	-31.0	24.8/4.2	-0.05
NCAR PCM1	0.82	0.8	-0.15	-0.58	-0.50	1.01	18.6	21.9/3.8	-0.08
MIROC-medres	-0.34	0.8	-0.08	0.06	0.12	2.28	1.0	13.0/1.3	-0.22
Observations	-1.98	2.5	0.03	1.28	-0.67	1.25	-134.0	16.4/3.5	0.43

^aAll characteristics are averaged for nine volcanic eruptions (except for NCEP/NCAR reanalysis data that cover only four volcanic eruptions since 1963 and U50, which is climatological mean) and for two winters (DJF) following volcanic eruptions as shown in Table 2. “Polar SLP” is the polar sea level pressure (SLP) anomaly averaged over a polar cap of 65°N–90°N. “Atl. SLP” is the maximum SLP anomaly at the Azores center. “T_{GL}” is the global surface air temperature anomaly. “T_{ES}” is the surface air temperature (SAT) anomaly averaged over Northern Eurasia (30°E–130°E; 45°N–70°N). “T_{AM}” is the SAT averaged over North America (120°W–60°W; 45°N–70°N). “T50” is the stratospheric temperature anomaly at 50 hPa averaged over the equatorial belt 0°–30°N. “H50” is the geopotential height anomaly at 50 hPa averaged over the polar cap 65°N–90°N. “U50/σ” are the climatologically mean zonal wind at 50 hPa averaged over the latitude belt 55°N–65°N and its standard deviation. “Niño3.4” is the Niño3.4 index, the sea surface temperature anomaly averaged over 170°W–120°W, 5°S–5°N. Statistically significant anomalies (with respect to climate variability) at the 90% confidence level are shown in bold italic.

anomaly values diagnosed for each model are given in the body of the table. Anomalies significant at the 90% confidence level are shown in bold italics. For the maps of the anomaly fields (Figures 2–7), any regions where the response is judged different from zero at least at the 90% confidence level are marked by hatching. The statistical significance was computed using a two-tailed t-test assuming each volcano and each model realization represent independent samples. Composites were computed using all nine volcanoes listed in Table 2 for all the models, except for NCAR PCM1 for which the analysis only includes the six post-1890 eruptions.

[22] It is worth mentioning that AO sensitivity to external forcing in the models is often presented using normalized empirical orthogonal functions. This can conceal large differences between the amplitude of the response in different models. To avoid this complication we specifically characterize the AO response in absolute quantities assuming that radiative forcing has comparable magnitude in all models and in the real world.

4.1. Regional Surface Temperature Response

[23] Figure 2 depicts surface air temperature anomaly composites for each of the models along with the comparable observed pattern. The observed pattern of temperature anomaly (Figure 2h) is consistent with the expectation that the AO is in an anomalously positive phase in the post-eruption periods [Thompson and Wallace, 1998]. Much of the observed warming reaches 90% confidence level in Northern Europe, Siberia, and eastern Asia. The cooling in the Middle East (which is another distinctive feature of the surface temperature anomalies in the positive AO phase) reaches -0.6 K but is not statistically significant. The observations do not show any warming in North America. The simulated cooling of about -0.8 K is not statistically significant but is qualitatively different from the warming anomaly that one would expect as a part of the positive AO phase pattern [Thompson and Wallace, 1998]. Observed warming in Central and South America, as well as along the west coast of North America is presumably to be attributed to a net sampling of positive ENSO phases in the

composite, rather than any volcanic effect. The composite SST anomaly in the equatorial East Pacific reaches about 0.5 K. None of the warming areas at low latitudes in the observed composite has statistical significance.

[24] In all the composites from the model simulations (Figures 2a–2g), the Southern Oscillation cycle is reasonably well averaged out. The SST anomaly in the Equatorial East Pacific in the model composites tends to be negative but does not exceed 0.2 K in magnitude (see Table 3, Niño3.4 index).

[25] Both GFDL (Figures 2a and 2b) and GISS (Figures 2c and 2d) models produce spatial patterns of winter warming over Eurasia that are in reasonable agreement with observations. The magnitude of the anomalies in the GISS model appears to be unrealistically small, however. The GFDL CM2.0 concentrates warming in central North Siberia, thus shifting it a bit excessively poleward. The GFDL CM2.1 correctly produces maximum warming in Europe and eastern Asia but underestimates the amplitude of warming in eastern Asia. The composite anomaly maps for both NCAR models display maximum warming very far to the north. The MIROC-medres model shows warming mostly over Europe. All models produce the observed cooling in the Middle East. The GFDL and NCAR models do not show significant warming over North America. Both GISS models tend to produce positive anomalies on the east coast of North America.

[26] The overall post-volcanic cooling of the surface expected from the reduction in solar radiation reaching the surface is found in all the models, particularly in the tropics and subtropics. By contrast with the fairly uniform cooling at low latitudes in the model results, the observed composite has a much more complicated appearance. Notably the observed post-volcano composite shows surface warming in the Eastern Pacific. As noted earlier, this is most reasonably interpreted as resulting from a limited sampling that leaves a significant net El Niño signal in the composite rather than an indication of the volcano-induced temperature anomaly. However, the models have more cooling than the observations in all the ocean basins suggesting that models might overestimate the volcanic impact on SST.

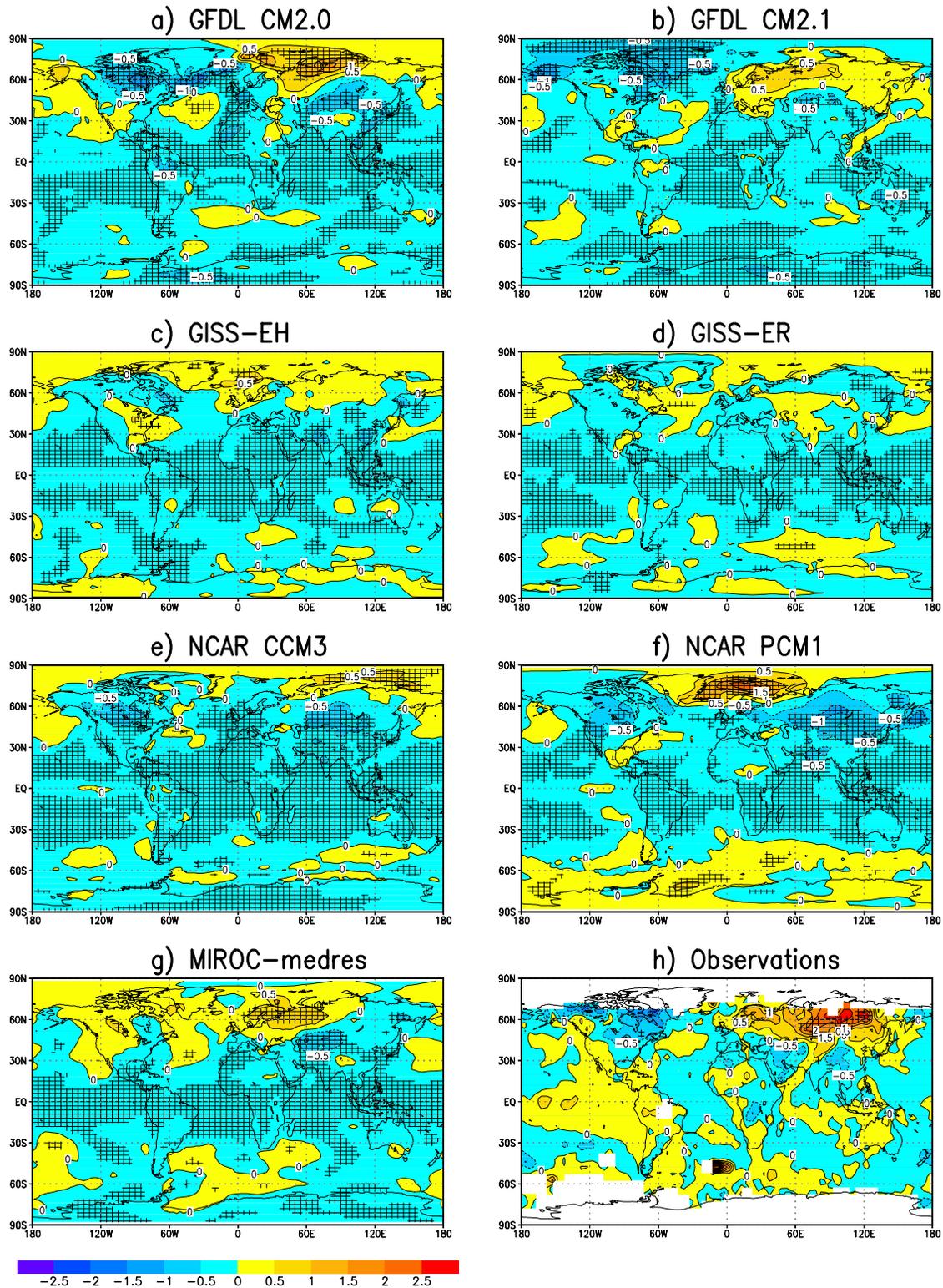


Figure 2. Surface winter (DJF) air temperature anomalies (K) composited for nine volcanic eruptions (see Table 2) and averaged for two seasons and all available ensemble members: (a–g) IPCC model simulations marked as in Table 1; (h) observations from HadCRUT2v data set. Hatching shows the areas with at least 90% confidence level calculated using a two-tailed local t-test.

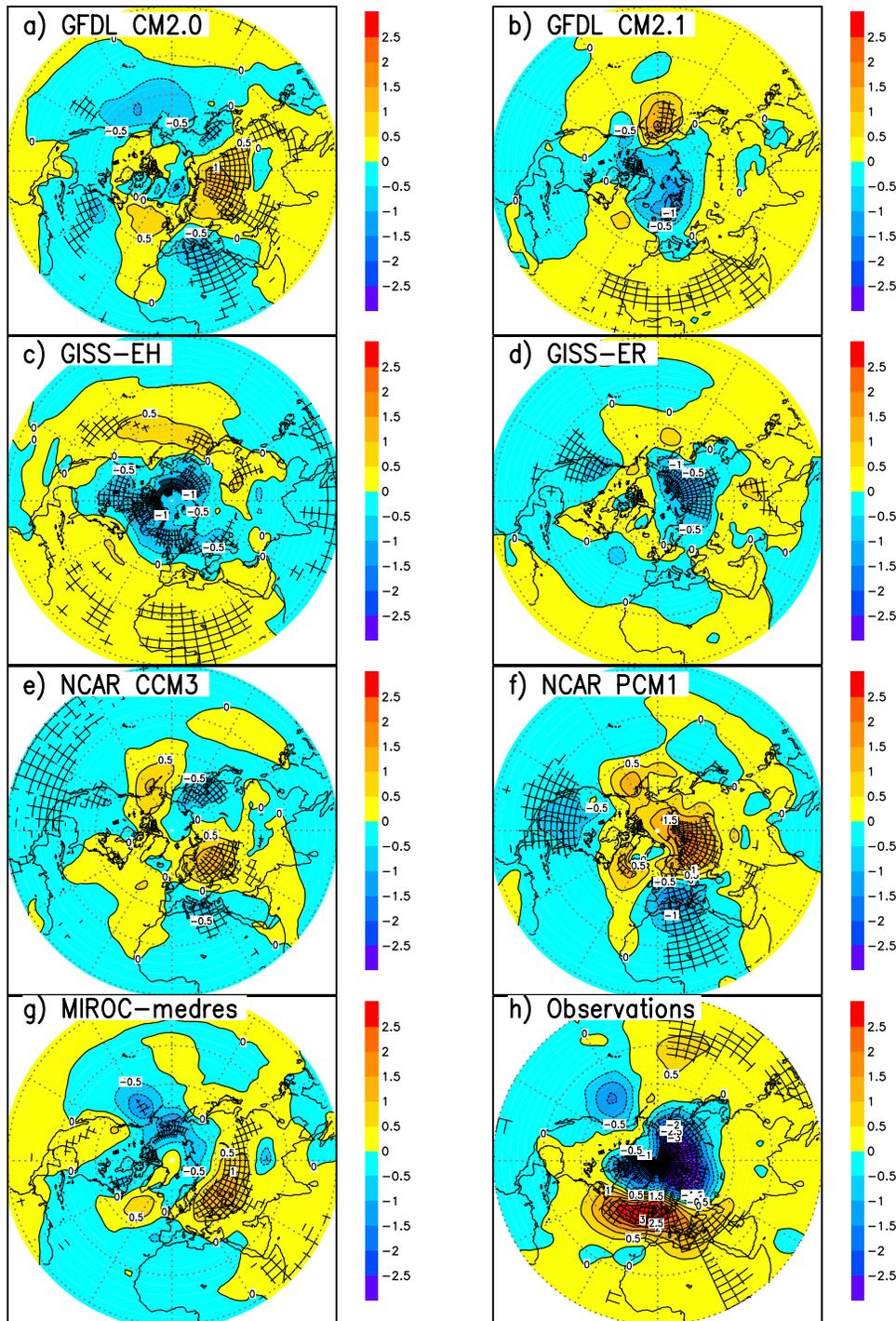


Figure 3. Sea level pressure anomalies (hPa) averaged for winter season (DJF) and composited for nine volcanic eruptions (Table 2) and averaged for two seasons and all available ensemble members: (a–g) IPCC model simulations marked as in Table 1; (h) observations from HadSLP1 data set. Hatching shows the areas with at least 90% confidence level calculated using two-tailed local t-test.

4.2. Sea Level Pressure Response

[27] The winter high-latitude warming seen in the observed post-volcano climate is consistent with anomalous circulation patterns featuring a strengthening of the tropospheric zonal wind and a poleward shift of storm tracks [Hurrell, 1995; Walter and Graf, 2005]. A composite of the

observed sea level pressure (SLP) anomalies for the post-volcanic periods is shown in Figure 3h along with the results from each of the models in Figures 3a–3g. The observed composite shows a strong low pressure anomaly centered near the North Pole. The high-latitude negative SLP anomaly is surrounded by a ring of positive SLP anomaly, but this is most pronounced over the Atlantic

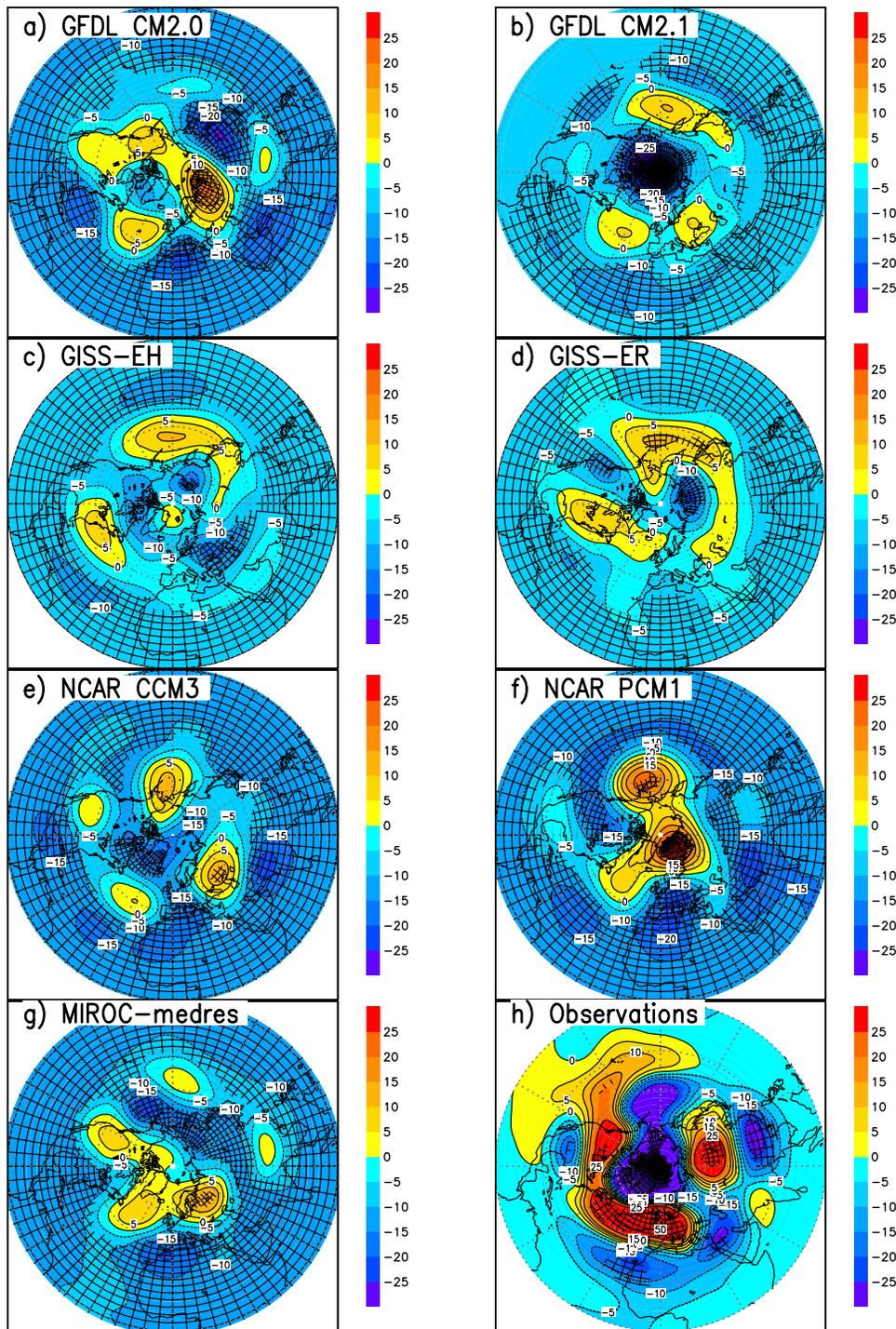


Figure 4. Geopotential height anomalies (m) at 200 hPa averaged for winter season (DJF) and composited for nine volcanic eruptions (Table 2) and averaged for two seasons and all available ensemble members: (a–g) IPCC model simulations marked as in Table 1; (h). NCEP/NCAR reanalysis data for period 1958–2000. Hatching shows the areas with at least 90% confidence level calculated using two-tailed local t-test.

sector and the Mediterranean region. The SLP anomaly over the Azores in the observed composite is +2.5 hPa. The strong meridional SLP gradient in the Atlantic sector drives westerly surface wind anomalies that help account for the corresponding warm surface anomalies in Northern Europe and Asia. The observed post-volcanic SLP anomalies in the

Pacific sector are much weaker. A negative SLP anomaly near the west coast of North America is caused presumably by a residual El Niño signal left in the observed composite. Alternatively, in the case of a strong polar vortex, *Perlwitz and Graf* [2001] showed that such an anomaly pattern can be produced by reflection of a zonal wave one disturbance

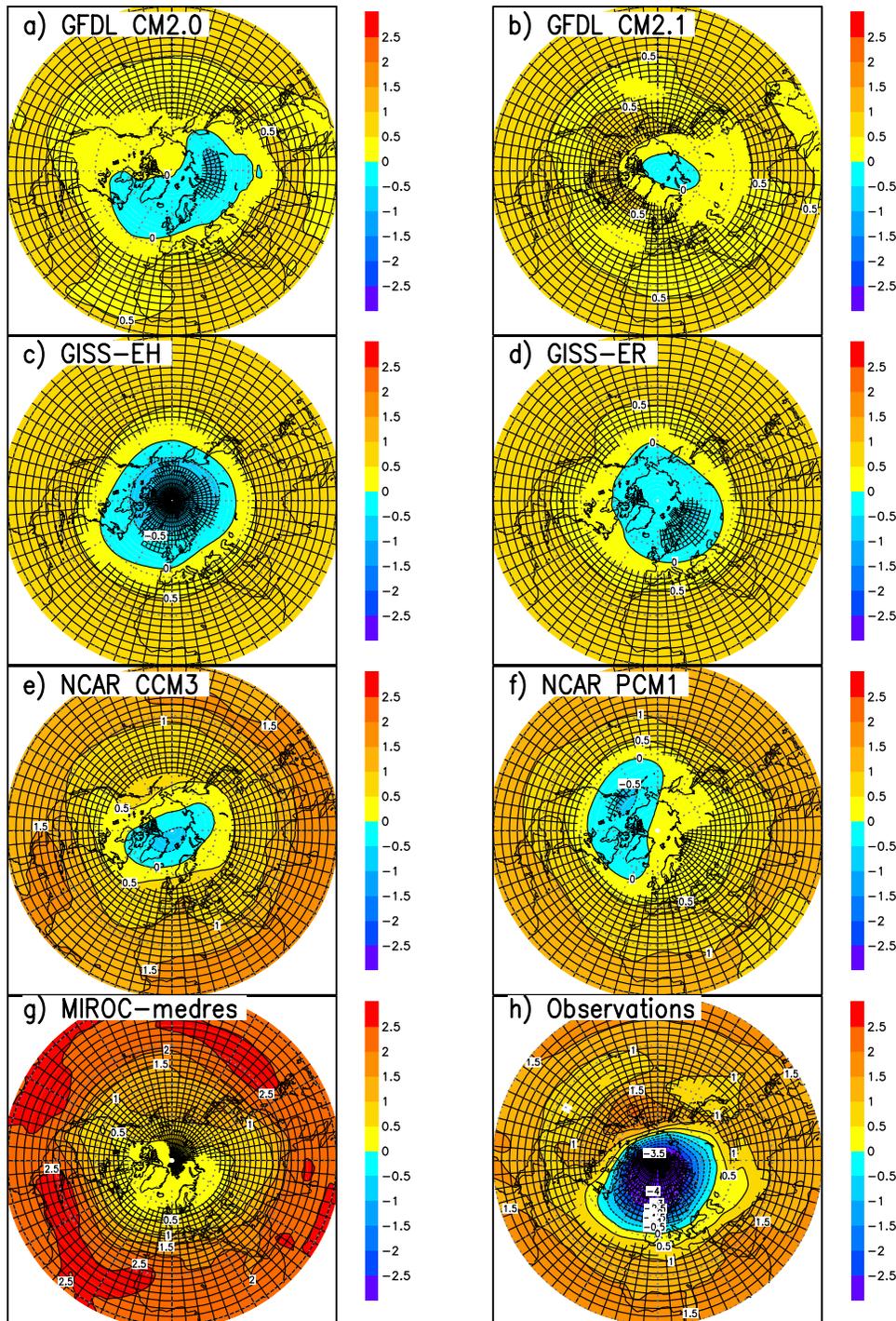


Figure 5. Atmospheric temperature anomalies (K) at 50 hPa averaged for winter season (DJF) and composited for nine volcanic eruptions (Table 2) and averaged for two seasons and all available ensemble members: (a–g) IPCC model simulations marked as in Table 1; (h) NCEP/NCAR reanalysis data for period 1958–2000. Hatching shows the areas with at least 90% confidence level calculated using two-tailed local t-test.

by negative wind shear in the mid-stratosphere [Perlwitz and Harnik, 2003].

[28] In most of the model SLP composites there is a similar basic pattern of low pressure over the pole surrounded by a ring of anomalously high pressure. However,

beyond that basic feature the models vary considerably in their SLP anomaly patterns. The GFDL and GISS model ensemble average results display the Atlantic dipole pattern of the observed sign, but weaker than in the observed composite. The Azores maxima in all the GFDL and GISS

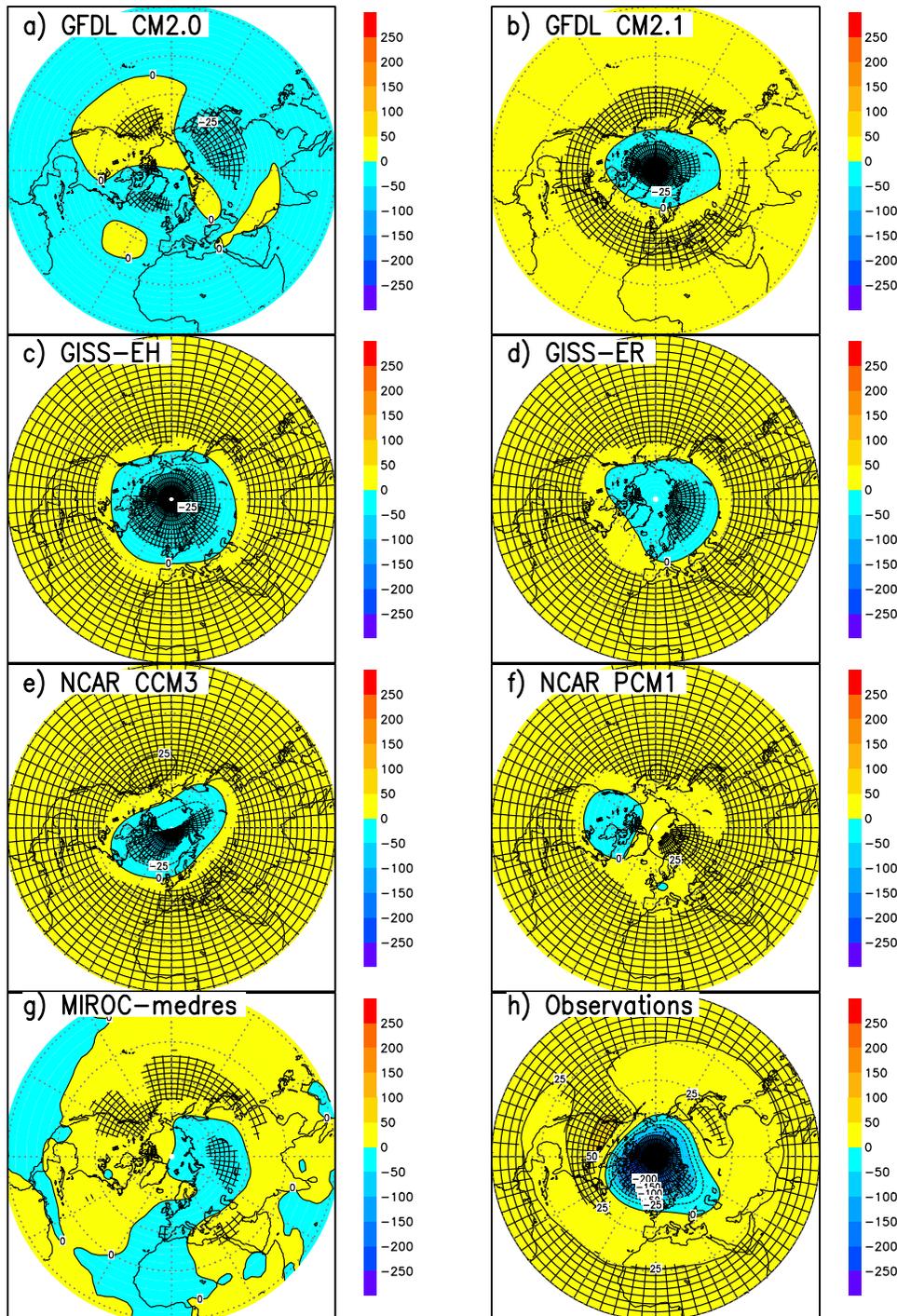


Figure 6. Geopotential height anomalies (m) at 50 hPa averaged for winter season (DJF) and composited for nine volcanic eruptions (Table 2) and averaged for two seasons and all available ensemble members: (a–g) IPCC model simulations marked as in Table 1; (h) NCEP/NCAR reanalysis data for period 1958–2000. Hatching shows the areas with at least 90% confidence level calculated using two-tailed local t-test.

models are about 0.5 hPa (Figures 3a–3d). The two NCAR models and the MIROC model have less clearly developed Atlantic sector dipole responses, and generally a noisier SLP anomaly composite. The NCAR PCM composite SLP anomaly does not show the polar minimum that is seen in observations and the other model results.

4.3. Tropospheric Jet and Storm Tracks

[29] Figure 4 shows the post-volcano composites of 200 hPa geopotential height anomalies for the seven models and observations. There were no global observations of stratospheric and upper tropospheric fields until the second half of the 20th century, and so the observed composite

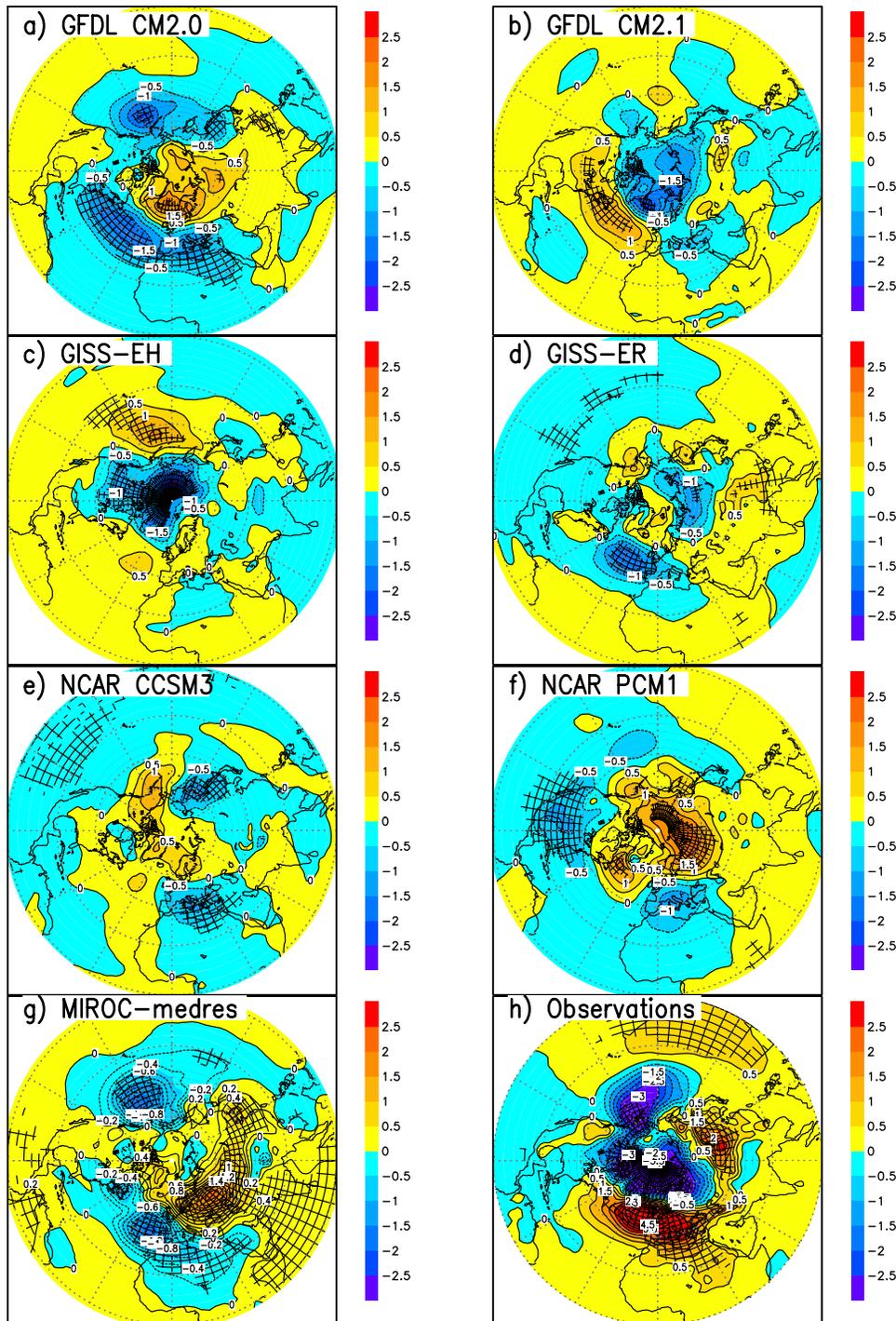


Figure 7. Sea level pressure anomalies (hPa) averaged for the winter season (DJF) and composited for four volcanic eruptions (Agung, Fuego, El Chichón, and Pinatubo, see Table 2) from the second half of the 20th century and averaged for two seasons and all available ensemble members: (a–g) IPCC model simulations marked as in Table 1; (h) NCEP/NCAR reanalysis data for period 1958–2000. Hatching shows the areas with at least 90% confidence level calculated using a two-tailed local t-test.

(Figure 4h) was computed using NCEP/NCAR reanalyses and for only the last 4 eruptions considered. We use the 200 hPa anomaly as a proxy to characterize the shift of the tropospheric jet assuming that this shift leads to displacement of the baroclinicity region in the upper troposphere and an associated shift of the storm track. The observed

200 hPa height anomalies in Figure 6h exceed 25 m over the North Atlantic, North America, and Eurasia. They show changes in the position of the tropospheric jet streams. The shift of tropospheric jets is accompanied by changes in the storm tracks that lead to increased precipitation and surface temperatures in North America and high-latitude Eurasia

[Hurrell, 1995]. The observed pattern shows that tropospheric jets move north in the North Atlantic sector, and over North America and Eurasia. By contrast in the Pacific sector they move equatorward.

[30] The model composite anomalies in 200 hPa heights vary enormously among the models considered here. The observed composite shows both a strong negative anomaly at the pole and positive anomalies in the Atlantic and European sectors. The GISS and GFDL model composites have a similar zonal-mean structure with negative anomalies at the pole and a positive in the zonal-mean anomaly in midlatitudes. However, their midlatitude anomalies are weaker than observed and often are out of phase with observations. The 200 hPa composite results for the NCAR models and the MIROC model are even further from agreement with the observations. The NCAR PCM1 even displays a mean meridional gradient reversed from that observed, with the model showing a positive height anomaly over the pole in the post-volcano winters.

4.4. Stratospheric Response

[31] The strongest direct radiative effects of volcanic aerosol loading are expected in the stratosphere, and it is likely that much of the tropospheric circulation response is caused by the dynamical stratospheric influence on the troposphere. The direct radiative effect of volcanic aerosols in the lower stratosphere in the layer of 30 hPa to 70 hPa acts to increase the equator-to-pole temperature contrast in the winter hemisphere and this leads to a stronger polar vortex. A stronger vortex will tend to reflect more of the quasi-stationary planetary waves forced in the troposphere. This leads to a reduced wave drag on the vortex and thus a positive feedback that should lead to further strengthening of the vortex.

[32] Figure 5 shows model and NCEP/NCAR reanalysis composites of 50 hPa temperature anomaly computed in the same manner as the 200 hPa geopotential heights in Figure 4 just discussed. The NCEP/NCAR data show an average warming of the equatorial region at 50 hPa of about 1.3K (see Table 3). (The four eruptions composited in this period include the very large Agung and El Chichón cases as well as the extremely large Pinatubo case, so one may expect the effects to be somewhat stronger than in the nine volcano composite adopted for the earlier results.)

[33] The model composites all show warming in the equatorial region at 50 hPa (see Table 3). The GFDL and GISS models produce temperature anomalies of 0.6–0.7 K. An independent analysis of stratospheric temperature responses for the recent eruptions of El Chichón and Pinatubo shows that the GFDL CM2.1 produces fairly realistic warming in the lower stratosphere in response to volcanic forcing [Ramaswamy *et al.*, 2006]. Stratospheric warming is substantially larger for the MIROC and NCAR models. The most likely explanation of these differences is in terms of the detailed treatment of the properties of the aerosols adopted in each model. The models exhibiting the most warming have likely larger absorption from stratospheric aerosols in the longwave and/or solar near-IR, both of which would lead to larger radiative heating rates in the equatorial lower stratosphere.

[34] The increased heating in NCAR models in comparison with GISS and GFDL models is most likely because

the aerosol optical depth of Ammann *et al.* [2003] used by NCAR is slightly larger than the one of Sato *et al.* [1993]. The aerosol radiative effect integrated over the whole spectrum is also sensitive to aerosol size distribution and refractive index. In the GISS and GFDL models stratospheric aerosol effective radius varies in time and space from about 0.6 μm to 0.1 μm for post-volcanic periods and decreases to 0.05 μm for background aerosols. The NCAR models use fixed effective radius of 0.42 μm not accounting for its spatial and temporal changes. The MIROC model assumes a fixed effective radius of 0.0695 μm that is more representative for background stratospheric aerosols. Stenchikov *et al.* [1998] pointed out that small aerosol particles for an aerosol layer with the same optical depth cause more near-IR absorption than large particles. This could be the reason why the MIROC model produces the largest temperature anomaly at 50 hPa in the tropics (T50, Table 3).

[35] Another mechanism that could affect the simulated temperature anomaly in the equatorial lower stratosphere could be increase of the cross-tropopause water vapor flux because of volcanically induced tropical tropopause warming [Joshi and Shine, 2003]. The extra water vapor in the stratosphere absorbs upward terrestrial and solar near-IR radiation, perturbing heating rates in the lower stratosphere. Although Joshi and Shine [2003] show that expected water vapor changes would be small and HALOE observations [Randel *et al.*, 2004] do not show a large effect for the post-Pinatubo period, it could be erroneously reproduced in the models and has to be thoroughly tested.

[36] All the models simulate the positive temperature anomaly in the lower equatorial stratosphere caused by radiative heating in the aerosol layer. In almost all cases there is net anomalous cooling over a region near the North Pole associated with strengthening of the polar vortex and negative anomalies of the 50 hPa geopotential height, except for the MIROC model, for which the composited 50 hPa temperature anomalies are positive even at the Pole, and for GFDL CM2.1, for which polar cooling is very weak. None of the models simulates as deep a polar cooling as found in the observed composite. The cooling observed near the pole is likely a dynamical consequence of the aerosol radiative perturbations, involving the effects of the mean conditions on the propagation of quasi-stationary planetary waves (see discussion section). Thus, while the differences in stratospheric equatorial warming among the models is most reasonably attributed to differences in the radiative effects, the differences at higher latitudes may reflect more subtle aspects of the way models represent the dynamics of the stratosphere. Even the GFDL CM2.1 model, which produced the best surface air temperature response, has trouble dynamically cooling the polar lower stratosphere.

[37] Figure 6 shows the post-volcanic winter composites of 50 hPa geopotential height from the models and observations. Once again the observed composite (Figure 6h) is computed using NCEP/NCAR data and only for the last four eruptions. Figure 5h shows that a very strong and statistically significant strengthening of the polar vortex occurs in post-volcanic winters. The 50 hPa geopotential height anomaly reaches -200 m near the North Pole.

[38] The model composites for the 50 hPa geopotential anomaly display a rather wide range of behavior, although

Table 4. Integrated Model Responses^a

Model Name	Polar SLP, hPa	Atl. SLP, hPa	T _{GL} , K	T _{ES} , K	T _{AM} , K	T50, K	H50, m	U50/σ, m/s	Niño3.4, K
GFDL CM2.0	0.62	1.6	-0.10	-0.32	0.07	0.57	13.9	19.2/3.2	-0.34
GFDL CM2.1	-1.06	1.4	-0.12	0.35	-0.58	0.67	-17.9	20.9/2.2	-0.19
GISS-EH	-1.15	0.6	-0.13	-0.17	0.07	0.99	-45.0	19.9/2.6	-0.09
GISS-ER	-0.22	0.2	-0.08	-0.14	-0.04	0.84	-9.5	19.0/2.3	-0.06
NCAR CCSM3	0.35	0.4	-0.08	-0.15	-0.20	1.77	-39.4	24.8/4.2	-0.04
NCAR PCM1	1.15	0.8	-0.12	-0.53	-0.49	1.30	15.4	21.9/3.8	0.02
MIROC-medres	0.22	1.6	-0.08	-0.23	0.08	2.74	-1.6	13.0/1.3	-0.15
Observations	-2.61	5.0	-0.03	1.56	-0.18	1.25	-134.0	16.4/3.5	0.22

^aSame as in Table 3, but all characteristics are averaged for four volcanic eruptions for the period since 1963, which is covered by the NCEP/NCAR reanalysis. The observed sea level pressure and surface air temperature indices are calculated using NCEP/NCAR reanalysis. Statistically significant anomalies (with respect to climate variability) at the 90% confidence level are shown in bold italic.

in each case there is at least a hint of the deepening of the polar low and consequent strengthening of the westerly vortex during the post-volcano winters. The models which simulate the most realistic winter surface warming over Northern Europe and Asia and the most realistic SLP anomaly patterns are those that also simulate a statistically significant strengthening of the polar vortex (Figures 6b–6d). The NCAR PCM1 and MIROC models, despite being forced by the strongest equatorial heating in the lower stratosphere, do not produce a substantial strengthening of the polar vortex. This indicates once more that nonlinearity of stratosphere-troposphere interaction (leading to positive feedback) plays an important role in the entire process.

[39] For completeness we also analyzed a reduced composite of four relatively strong recent eruptions in the second part of the 20th century for the period that is covered by the NCEP/NCAR reanalysis, calculated in the same way as composites for the NCEP/NCAR reanalysis data in Figures 4–6. They appear to be qualitatively similar to the 9-volcano composites. The GISS-EH and GFDL CM2.1 4-volcano patterns look closer to observations, however, those of GISS-ER and GFDL CM2.0 deteriorated. Both NCAR and MIROC models fail to reproduce the observed patterns as in the 9-volcano analysis. This point is well illustrated by the sea level pressure composite anomalies shown in Figure 7 and in Table 4, which shows the same indices as in Table 3 but calculated for 4-volcano composites. In Table 4 the GFDL CM2.1 and GISS-EH polar SLP anomalies increased in magnitude, exceeding 1 hPa, but still remained significantly lower than observed ones. All models in Table 4 show significant global surface cooling of about the same magnitude as in Table 3 and about 20% stronger lower stratospheric warming responding to the stronger composite-average radiative forcing than in the 9-volcano composites. Winter warming over Eurasia, as measured by our index TES, is produced by only GFDL CM2.1. However, strengthening of the polar vortex is better captured by GISS-EH (H50 reaches -45 m). NCAR CCSM3 also produced a very strong geopotential height anomaly at 50 hPa reaching -39.4 m. Overall, 4- and 9-volcano analyses appear to be fairly consistent, showing that the simulated AO responses and associated stratospheric annular perturbations are significantly weaker than observed.

5. Summary and Discussion

[40] This paper reports on a diagnosis of the effects of volcanic eruptions on the simulated climate in a number of

the models participating in the IPCC AR4 model intercomparison. The analysis of anomalies in the post-volcanic periods in the “historical” IPCC runs provides a test of how the different models respond to a global radiative perturbation. These model experiments had similar designs in terms of how the volcanic aerosol effects were imposed. Quantitative offline calculations of the radiative perturbations caused by the imposed aerosol in each model are not available. However, the effect of the aerosol in reflected solar radiation can be diagnosed approximately from the TOA outgoing solar radiation in the actual model simulations. It is clear that the different models have a volcanic aerosol specification with fairly similar effects on the solar radiation, although there are some consistent differences among the models in this respect. Notably the post-volcanic aerosol in the MIROC model seems to produce a substantially smaller increase in the global albedo than that used in the other models.

[41] The post-volcano global-mean surface cooling expected from the increased global albedo is seen in each model analyzed. Averaged over the two boreal winters after each of the nine eruptions and averaged over all realizations performed, the global-mean surface temperature anomalies in the models varied from -0.06 K to -0.17 K (Table 3). While these values seem reasonable, it is difficult to make any further inference about how realistic each model is in this respect. The observed composite actually has a very weak anomalous warming of 0.02 K when averaged for the same set of eruptions. The observed value no doubt has substantial temporal sampling errors, and the observations for the earliest eruptions considered may suffer from inadequate geographical sampling or other data quality issues. It is noteworthy in this respect that the global-mean surface temperature record used in this study surprisingly shows no global cooling following the very large 1883 Krakatau eruption [Jones *et al.*, 2003]. However, the surface air temperature reconstructed by Hansen and Lebedeff [1988] shows a more sizable cooling effect of Krakatau, which indicates the level of uncertainty in the observations themselves, especially for earlier volcanic events.

[42] In the two years following major eruptions, the NH winter tropospheric circulation has been observed typically to display features characteristic of an anomalously positive AO index situation. This has a zonal-mean expression with low-pressure at high latitudes and a ring of anomalously high pressure in the midlatitudes. This basic zonal-mean pattern is modulated by a very strong regional structure with an intensified high pressure anomaly over the North Atlan-

tic and Mediterranean sectors. Consistent with this are shifts in the Atlantic storm track and an increased flow of warm air to Northern Europe and Asia, where anomalously high winter surface temperatures are observed.

[43] The models considered here display only limited success in reproducing these observed tropospheric post-volcano circulation and thermal anomalies. The GFDL, GISS and MIROC models all show a tendency for the model anomalies to have roughly the same zonal-mean signature as seen in observations, with lower surface pressure near the pole, but the zonal structure of the observed anomalies is not well reproduced by any of the models. Notably the strong anomalous high surface pressure over the Atlantic sector seen in observations is not adequately simulated in any of the models. The NCAR CCSM3 and PCM1 models stand out with the least realistic simulated anomalies at the surface. The results for these two models show no clear signal even of the zonal-mean structure of the expected positive AO anomaly index flow, and also display anomalously low winter temperatures over most of Northern Europe and Asia where observations have shown a consistent tendency for warming after almost all individual eruptions [Robock and Mao, 1992]. The most realistic patterns of the high-latitude surface warming appear to be produced by the GFDL CM2.1 (see also Tables 3 and 4).

[44] The most direct effect of the volcanic aerosol loading should be in the stratosphere. It is known that post-eruption periods are characterized by anomalously warm tropical lower stratospheric conditions and, in the NH winter, an anomalously cold and intense polar vortex. The tropical temperature anomalies at 50 hPa are a direct response to the enhanced absorption of terrestrial IR and solar near-IR radiation by the aerosols. The high-latitude winter perturbations at 50 hPa are a dynamical response to the strengthening of the polar vortex. The models all produce a substantial warming of the tropical stratosphere in the post-volcano periods, as expected, although the MIROC model stands out as having substantially greater warming than the other models or the observations. The average winter polar vortex perturbations in the post-volcano composites vary a great deal amongst the models, but except for MIROC and NCAR PCM1, the models simulate a colder and stronger polar vortex than normal.

[45] In this comparison, the ensemble average response from the models for two NH winters (nine volcanoes and at least three different realizations) is compared to the single natural realization for the same nine volcanoes and two NH winters. In both models and observations, sampling is an important issue. Due to the imperfect sampling of these noisy fields, one should not expect perfect agreement between the models and observations. Furthermore the increased averaging of the model response as compared with the observations, should produce a more statistically significant signal but with lower variability. But because the strongest simulated responses are 3–10 times weaker than observed (compare, for example, Figures 3b and 3h for sea level pressure anomaly) we conclude that the model AO sensitivity is not as strong as in observations.

[46] It is interesting that the sea level pressure principal component analysis of AO sensitivity to volcanic forcing conducted by Miller *et al.* [2006], using a subset of the five strongest eruptions from Table 2, gives similar results in

terms of relative model behavior and the weakness of model responses in comparison with observations. This supports the credibility and robustness of our conclusions. Miller *et al.* [2006] also showed that, despite observations, neither the GISS nor the GFDL models produced a significant AO trend in the 19th to 20th centuries. Only in the 21st century, when anthropogenic forcing strengthens, did the GISS and GFDL models produce a positive AO trend.

[47] Determining why the model AO response is too weak would require additional analysis. However, one possibility is that the models may simply not have sufficiently fine resolution and sufficiently deep model domains to adequately treat stratospheric dynamics and the stratosphere-troposphere dynamical interactions. All of the models considered here have rather coarse vertical resolution in the stratosphere, and the GFDL, NCAR and MIROC models have model tops at or below the stratopause. The GISS models do have the model top at 0.1 hPa, but these models have the coarsest vertical and horizontal resolution of those considered here (and coarse relative to most of the models in the IPCC intercomparison). There is increasing evidence of the synoptic-scale process contribution to AO development [Benedict *et al.*, 2004; Frantze *et al.*, 2004; Song and Robinson, 2004; Wittman *et al.*, 2004]. Underestimation of synoptic activity in the current GCMs might be one of the reasons for the low AO sensitivity. However, it is not obvious from the AR4 model results, because higher spatial resolution (like in NCAR CCSM3) does not lead automatically to higher AO sensitivity [Miller *et al.*, 2006].

[48] One consequence of the inadequate treatment of stratospheric dynamics in the models may be a mean climate characterized by an unrealistically intense polar vortex. If the winter vortex is too strong it may be unrealistically resistant to penetration by planetary waves, and thus much too stable. All models (except MIROC-medres) in Table 3 have stronger zonal winds than observed and all of them (except both NCAR models) have weaker variability of zonal winds than observed. This problem could be expected to weaken any wave feedback in the models, and possibly prevent the propagation of stratospheric signals into the troposphere. The NCAR CCSM3, for example, has the strongest climatological zonal wind at 50 hPa (Table 3). Therefore although it produces reasonably strong zonal-mean stratospheric anomalies in response to volcanic forcing they do not lead to a realistic propagation of the effect into the troposphere. Castanheira and Graf [2003] and Walter and Graf [2005] showed that a 20 m/s zonal mean wind at 50 hPa at the polar circle is a good estimate for the transition of vertical wave propagation regimes. If the winds are above that limit a typical positive AO pattern emerges with all the effects of mild winters over Eurasia. All the models (except MIROC-medres) are close to or above this threshold in the climatological mean. Hence, the strengthening of the zonal winds by volcanic aerosol effects may not significantly change the frequency of positive AO any further and this very possibly may be an additional reason for the models underestimating winter warming.

[49] Figure 6 also shows that the simulated 50 hPa geopotential height anomalies in the equatorial region are lower than in observations. This is because of the larger sampling size; El Niño was averaged out in simulations but has

sizable amplitude in the observation composite (see Figure 1 and Niño3.4 index in Tables 3 and 4). The tropospheric cooling in the simulation composites causes shrinking of the tropospheric column in the tropics (Figure 4) that reduces the increase of geopotential height at 50 hPa. However, it is known that El Niño affects North American winter surface air temperature and only weakly affects temperature over Eurasia [Yang and Schlesinger, 2002]. L'Heureux and Thompson [2006] showed that the El Niño high-latitude response projects onto the Southern Annular Mode but not to the Northern Annular Mode. Therefore we assumed that the El Niño signal will not dominate in the observed AO responses. However, Rind et al. [2005] reported a positive North Atlantic Oscillation response to an El Niño-type SST anomaly in an older version of the GISS GCM.

[50] It will also be interesting to compare the responses of the models to volcanic aerosol loading documented here with other aspects of the model behavior being analyzed in the IPCC AR4 projects. As noted above, most earlier hindcasts of 20th century climate as well as current IPCC AR4 runs [Miller et al., 2006; Knutson et al., 2006] do not reproduce the observed trends over recent decades in the AO component of the circulation, and thus do not capture the intensification of warming trends that has been observed over Northern Europe and Asia. There are various possible explanations for this discrepancy, but it is interesting to speculate that it could indicate that the models employed may have a basic inadequacy that does not allow a sufficiently strong AO response to large-scale forcing, and that this inadequacy could also be reflected in the simulated response to volcanic aerosol loading.

[51] **Acknowledgments.** We thank the international modeling groups for providing their data for analysis, P. Jones, D. Shindell, and J. Perlwitz for comments, the Program for Climate Model Diagnosis and Intercomparison for collecting and archiving the model data, the JSC/CLIVAR Working Group on Coupled Modeling and their Coupled Model Intercomparison Project and Climate Simulation Panel for organizing the model data analysis activity, and the IPCC WG1 TSU for technical support. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy. G.S. and A.R. supported by NSF grants ATM-0313592 and ATM-0351280 and NASA grant NNG05GB06G. A.R. supported by the Chaire du Développement Durable de l'École Polytechnique, Paris, France. This research was also supported in part by the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) through its sponsorship of the International Pacific Research Center.

References

- Adams, J. B., M. E. Mann, and C. M. Ammann (2003), Proxy evidence for an El Niño-like response to volcanic forcing, *Nature*, *426*, 274–278.
- Ammann, C., G. Meehl, W. Washington, and C. Zender (2003), A monthly and latitudinally varying forcing data set in simulations of 20th century climate, *Geophys. Res. Lett.*, *30*(12), 1657, doi:10.1029/2003GL018875.
- Andronova, N., E. Rozanov, F. Yang, M. Schlesinger, and G. Stenchikov (1999), Radiative forcing by volcanic aerosols from 1850 through 1994, *J. Geophys. Res.*, *104*, 16,807–16,826.
- Basnett, T., and D. Parker (1997), Development of the global mean sea level pressure data set GMSLP2, *Clim. Res. Tech. Note 79*, Hadley Cent., Met Office, Exeter, U.K.
- Benedict, J., S. Lee, and S. Feldstein (2004), Synoptic view of the North Atlantic Oscillation, *J. Atmos. Sci.*, *61*, 121–144.
- Black, R. (2002), Stratospheric forcing of surface climate in the arctic oscillation, *J. Clim.*, *15*, 268–277.
- Black, R. X., and B. A. McDaniel (2004), Diagnostic case studies of the northern annular mode, *J. Clim.*, *17*, 3990–4004.
- Broccoli, A., K. Dixon, T. Delworth, T. Knutson, and R. Stouffer (2003), Twentieth-century temperature and precipitation trends in ensemble climate simulations including natural and anthropogenic forcing, *J. Geophys. Res.*, *108*(D24), 4798, doi:10.1029/2003JD003812.
- Castanheira, J. M., and H.-F. Graf (2003), North Pacific–North Atlantic relationships under stratospheric control?, *J. Geophys. Res.*, *108*(D1), 4036, doi:10.1029/2002JD002754.
- Delworth, T. L., et al. (2006), GFDL's CM2 global coupled climate models, Part 1: Formulation and simulation characteristics, *J. Clim.*, *19*(5), 643–674.
- Frantze, C., S. Lee, and S. Feldstein (2004), Is the North Atlantic Oscillation a breaking wave?, *J. Atmos. Sci.*, *61*, 145–160.
- Gillett, N. (2005), Northern Hemisphere circulation, *Nature*, *437*, doi:10.1038/437496a, 496.
- Gnanadesikan, A., et al. (2006), GFDL's CM2 global coupled climate models, Part 2: The baseline ocean simulation, *J. Clim.*, *19*(5), 675–697.
- Graf, H.-F., and C. Timmreck (2001), Aerosol radiative forcing of the Laacher See volcano eruption, *J. Geophys. Res.*, *106*, 14,747–14,756.
- Graf, H.-F., I. Kirchner, A. Robock, and I. Schult (1993), Pinatubo eruption winter climate effects: Model versus observation, *Clim. Dyn.*, *9*, 81–93.
- Groisman, P. Y. (1992), Possible regional climate consequences of the Pinatubo eruption: An empirical approach, *Geophys. Res. Lett.*, *19*, 1603–1606.
- Hansen, J., and S. Lebedeff (1988), Global surface air temperatures: Update through 1987, *Geophys. Res. Lett.*, *15*, 323–326.
- Hansen, J. E., et al. (2002), Climate forcing in Goddard Institute for Space Studies S12000 simulations, *J. Geophys. Res.*, *107*(D18), 4347, doi:10.1029/2001JD001143.
- Haynes, P. H., M. E. McIntyre, T. G. Shepherd, C. J. Marks, and K. P. Shine (1991), On the “downward” control of extratropical diabatic circulations by eddy induced mean zonal forces, *J. Atmos. Sci.*, *48*, 651–680.
- Humphreys, W. J. (1913), Volcanic dust and other factors in the production of climatic changes, and their possible relation to ice ages, *J. Franklin Inst.*, 131–172.
- Humphreys, W. J. (1940), *Physics of the Air*, 676 pp., Dover, Mineola, N. Y.
- Hurrell, J. W. (1995), Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation, *Science*, *269*, 676–679.
- Jones, P. D., and A. Mobley (2003), Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001, *J. Clim.*, *16*, 206–223.
- Jones, P. D., T. J. Osborn, K. R. Briffa, C. K. Folland, E. B. Horton, L. V. Alexander, D. E. Parker, and N. A. Rayner (2001), Adjusting for sampling density in grid box land and ocean surface temperature time series, *J. Geophys. Res.*, *106*(D4), 3371–3380.
- Jones, P. D., A. Moberg, T. J. Osborn, and K. R. Briffa (2003), Surface climate responses to explosive volcanic eruptions seen in long European temperature records and mid-to-high latitude tree-ring density around the Northern Hemisphere, in *Volcanism and the Earth's Atmosphere*, *Geophys. Monogr. Ser.*, vol. 139, edited by A. Robock and C. Oppenheimer, pp. 239–254, AGU, Washington, D. C.
- Joshi, M. M., and K. P. Shine (2003), A GCM study of volcanic eruptions as a cause of increased stratospheric water vapor, *J. Clim.*, *16*, 3525–3534.
- Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, D. L. Williamson, and P. J. Rasch (1998), The National Center for Atmospheric Research Community Climate Model: CCM3, *J. Clim.*, *11*, 1131–1149.
- Kirchner, I., G. L. Stenchikov, H.-F. Graf, A. Robock, and J. C. Antuña (1999), Climate model simulation of winter warming and summer cooling following the 1991 Mount Pinatubo volcanic eruption, *J. Geophys. Res.*, *104*, 19,039–19,055.
- Kistler, R., et al. (2001), The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, *Bull. Am. Meteorol. Soc.*, *82*, 247–267.
- Knutson, T. R., T. L. Delworth, K. W. Dixon, I. M. Held, J. Lu, D. V. Ramaswamy, D. Schwarzkopf, G. Stenchikov, and R. J. Stouffer (2006), Assessment of twentieth-century regional surface temperature trends using the GFDL CM2 coupled models, *J. Clim.*, in press.
- Kodera, K. (1994), Influence of volcanic eruptions on the troposphere through stratospheric dynamical processes in the Northern Hemisphere winter, *J. Geophys. Res.*, *99*, 1273–1282.
- Kodera, K., and Y. Kuroda (2000a), Tropospheric and stratospheric aspects of the arctic oscillation, *Geophys. Res. Lett.*, *27*, 3349–3352.
- Kodera, K., and Y. Kuroda (2000b), A mechanistic model study of slowly propagating coupled stratosphere-troposphere variability, *J. Geophys. Res.*, *105*, 12,361–12,370.
- Labitzke, K. (1994), Stratospheric temperature changes after the Pinatubo eruption, *J. Atmos. Terr. Phys.*, *56*, 1027–1034.
- Labitzke, K., and M. P. McCormick (1992), Stratospheric temperature increases due to Pinatubo aerosol, *Geophys. Res. Lett.*, *19*, 207–210.
- Labitzke, K., B. Naujokat, and M. P. McCormick (1983), Temperature effects on the stratosphere of the April 4, 1982 eruption of El Chichón, Mexico, *Geophys. Res. Lett.*, *10*, 24–26.
- L'Heureux, L., and D. Thompson (2006), Observed relationships between the El-Niño/Southern Oscillation and the extratropical zonal-mean circulation, *J. Clim.*, *19*(2), 276–287, doi:10.1175/JCLI3617.1.

- Limpasuvan, V., D. J. W. Thompson, and D. L. Hartmann (2004), The life cycle of the Northern Hemisphere sudden stratospheric warmings, *J. Clim.*, *17*, 2584–2596.
- Mann, M. E., M. A. Cane, S. E. Zebiak, and A. Clement (2005), Volcanic and solar forcing of the tropical Pacific over the past 1000 years, *J. Clim.*, *18*, 447–456.
- Miller, R. L., G. A. Schmidt, and D. Shindell (2006), Forced variations of annular modes in the 20th century IPCC AR4 simulations, *J. Geophys. Res.*, doi:10.1029/2005JD006323, in press.
- Mitchell, J. M., Jr. (1961), Recent secular changes of the global temperature, *Ann. N. Y. Acad. Sci.*, *95*, 235–250.
- Newhall, C. G., and S. Self (1982), The Volcanic Explosivity Index (VEI): An estimate of explosive magnitude for historical volcanism, *J. Geophys. Res.*, *87*, 1231–1238.
- Ohhashi, Y., and K. Yamazaki (1999), Variability of the Eurasian pattern and its interpretation by wave activity flux, *J. Meteorol. Soc. Jpn.*, *77*, 495–511.
- Oman, L., A. Robock, G. Stenchikov, G. A. Schmidt, and R. Ruedy (2005), Climatic response to high latitude volcanic eruptions, *J. Geophys. Res.*, *110*, D13103, doi:10.1029/2004JD005487.
- Osborn, T. J. (2004), Simulating the winter North Atlantic Oscillation: The roles of internal variability and greenhouse gas forcing, *Clim. Dyn.*, *22*, 605–623, doi:10.1007/s00382-004-0405-1.
- Ostermeier, G. M., and J. M. Wallace (2003), Trends in the North Atlantic Oscillation—Northern Hemisphere annular mode during the twentieth century, *J. Clim.*, *16*, 336–341.
- Parker, D. E., and J. K. L. Brownscombe (1983), Stratospheric warming following the El Chichón volcanic eruption, *Nature*, *301*, 406–408.
- Parker, D. E., C. K. Folland, and M. Jackson (1995), Marine surface temperature: Observed variations and data requirements, *Clim. Change*, *31*, 559–600.
- Perlwitz, J., and H.-F. Graf (1995), The statistical connection between tropospheric and stratospheric circulation of the Northern Hemisphere in winter, *J. Clim.*, *8*, 2281–2295.
- Perlwitz, J., and H.-F. Graf (2001), Troposphere-stratosphere dynamic coupling under strong and weak polar vortex conditions, *Geophys. Res. Lett.*, *28*, 271–274.
- Perlwitz, J., and N. Harnik (2003), Observational evidence of a stratospheric influence on the troposphere by planetary wave reflection, *J. Clim.*, *16*, 3011–3026.
- Quiroz, R. S. (1983), The isolation of stratospheric temperature change due to the El Chichón volcanic eruption from nonvolcanic signals, *J. Geophys. Res.*, *88*, 6773–6780.
- Ramachandran, S., V. Ramaswamy, G. L. Stenchikov, and A. Robock (2000), Radiative impact of the Mount Pinatubo volcanic eruption: Lower stratospheric response, *J. Geophys. Res.*, *105*, 24,409–24,429.
- Ramaswamy, V., M. D. Schwarzkopf, W. Randel, B. Santer, B. J. Soden, and G. Stenchikov (2006), Anthropogenic and natural influences in the evolution of lower stratospheric cooling, *Science*, *311*, 1138–1141.
- Randel, W. J., F. Wu, S. Oltmans, K. Rosenlof, and G. Nedoluha (2004), Interannual changes of stratospheric water vapor and correlations with tropical tropopause temperatures, *J. Atmos. Sci.*, *61*, 2133–2148.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*(D14), 4407, doi:10.1029/2002JD002670.
- Rind, D., J. Perlwitz, and P. Lonergan (2005), AO/NAO response to climate change: 1. Respective influences of stratospheric and tropospheric climate changes, *J. Geophys. Res.*, *110*, D12107, doi:10.1029/2004JD005103.
- Robock, A. (2000), Volcanic eruptions and climate, *Rev. Geophys.*, *38*, 191–219.
- Robock, A. (2003), Introduction: Mount Pinatubo as a test of climate feedback mechanisms, in *Volcanism and the Earth's Atmosphere*, *Geophys. Monogr. Ser.*, vol. 139, edited by A. Robock and C. Oppenheimer, pp. 1–8, AGU, Washington, D. C.
- Robock, A., and J. Mao (1992), Winter warming from large volcanic eruptions, *Geophys. Res. Lett.*, *19*, 2405–2408.
- Robock, A., and J. Mao (1995), The volcanic signal in surface temperature observations, *J. Clim.*, *8*, 1086–1103.
- Sato, M., J. Hansen, M. P. McCormick, and J. Pollack (1993), Stratospheric aerosol optical depth, 1850–1990, *J. Geophys. Res.*, *98*, 22,987–22,994.
- Scaife, A., J. Knight, G. Vallis, and C. Folland (2005), A stratospheric influence on the winter NAO and North Atlantic surface climate, *Geophys. Res. Lett.*, *32*, L18715, doi:10.1029/2005GL023226.
- Shindell, D. T., G. A. Schmidt, R. L. Miller, and D. Rind (2001), Northern Hemisphere winter climate response to greenhouse gas, ozone, solar, and volcanic forcing, *J. Geophys. Res.*, *106*, 7193–7210.
- Shindell, D. T., G. A. Schmidt, M. Mann, and G. Faluvegi (2004), Dynamic winter climate response to large tropical volcanic eruptions since 1600, *J. Geophys. Res.*, *109*, D05104, doi:10.1029/2003JD004151.
- Song, Y., and W. A. Robinson (2004), Dynamical mechanisms for stratospheric influences on the troposphere, *J. Atmos. Sci.*, *61*, 1711–1725.
- Stenchikov, G., I. Kirchner, A. Robock, H.-F. Graf, J. C. Antuña, R. G. Grainger, A. Lambert, and L. Thomason (1998), Radiative forcing from the 1991 Mount Pinatubo volcanic eruption, *J. Geophys. Res.*, *103*, 13,837–13,857.
- Stenchikov, G., A. Robock, V. Ramaswamy, M. D. Schwarzkopf, K. Hamilton, and S. Ramachandran (2002), Arctic Oscillation response to the 1991 Mount Pinatubo eruption: Effects of volcanic aerosols and ozone depletion, *J. Geophys. Res.*, *107*(D24), 4803, doi:10.1029/2002JD002090.
- Stenchikov, G., K. Hamilton, A. Robock, V. Ramaswamy, and M. D. Schwarzkopf (2004), Arctic Oscillation response to the 1991 Pinatubo Eruption in the SKYHI GCM with a realistic Quasi-Biennial Oscillation, *J. Geophys. Res.*, *109*, D03112, doi:10.1029/2003JD003699.
- Thompson, D. W. J., and J. M. Wallace (1998), The Arctic Oscillation signature in the wintertime geopotential height and temperature fields, *Geophys. Res. Lett.*, *25*, 1297–1300.
- Walter, K., and H.-F. Graf (2005), The North Atlantic variability structure, storm tracks, and precipitation depending on the polar vortex strength, *Atmos. Chem. Phys.*, *5*, 1–10.
- Wittman, M. A. H., L. M. Polvani, R. K. Scott, and A. J. Charlton (2004), Stratospheric influence on baroclinic lifecycles and its connection to the Arctic Oscillation, *Geophys. Res. Lett.*, *31*, L16113, doi:10.1029/2004GL020503.
- Yang, F., and M. Schlesinger (2002), On the surface and atmospheric temperature changes following the 1991 Pinatubo volcanic eruption: A GCM study, *J. Geophys. Res.*, *107*(D8), 4073, doi:10.1029/2001JD000373.
- Yukimoto, S., et al. (2001), The new Meteorological Research Institute coupled GCM (MRI-CGCM2): Model climate and variability, *Pap. Meteorol. Geophys.*, *51*, 47–88.

H.-F. Graf, Department of Geography, University of Cambridge, Cambridge CB2 3EN, U.K.

K. Hamilton, International Pacific Research Center, University of Hawaii, Honolulu, HI 96822, USA.

V. Ramaswamy and R. J. Stouffer, NOAA Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, NJ 08542, USA.

A. Robock and G. Stenchikov, Department of Environmental Sciences, Rutgers University, 14 College Farm Rd., New Brunswick, NJ 08901, USA. (gera@envsci.rutgers.edu)

B. Santer, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA.